Data processing in quantum information theory

A.E. Allahverdyan, D.B. Saakian

Yerevan Physics Institute

Alikhanian Brothers St.2, Yerevan 375036, Armenia

Abstract

The strengthened data processing inequality have been proved. The general theory have been illustrated on the simple example.

Quantum information theory is a new field with potential applications for the conceptual foundation of quantum mechanics. It appears to be the basis for a proper understanding of the emerging fields of quantum computation, communication and cryptography [1-4]. Quantum information theory concerned with quantum bits (qubits) rather than bits. Qubits can exist in superposition or entanglement states with other qubits, a notion completely inaccessible for classical mechanics. More general, quantum information theory contains two distinct types of problem. The first type describes transmission of classical information through a quantum channel (the channel can be noisy or noiseless). In such scheme bits encoded as some quantum states and only this states or its tensor products are transmitted. In the second case arbitrary superposition of this states or entanglement states are transmitted. In the first case the problems can be solved by methods of classical information theory, but in the second case new physical representations are needed.

Mutual information is the most important ingredient of information theory. In classical theory this value was introduced by C.Shannon [9]. The mutual information between two ensembles of random variables X, Y (for example this ensembles can be input and output for a noisy channel)

$$I(X,Y) = H(Y) - H(Y/X), \tag{1}$$

is the decrease of the entropy of X due to the knowledge about Y, and conversely with interchanging X and Y. Here $H(Y)$ and $H(Y/X)$ are Shannon entropy and mutual entropy [9].

Mutual information in the quantum case must take into account the specific character of the quantum information as it is described above. The first reasonable definition of this quantity was introduced by B.Schumacher and M.P.Nielsen [2]. Suppose a quantum system with density matrix

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad \sum_i p_i = 1. \tag{2}$$

We only assume that $\langle\psi_i|\psi_i\rangle = 1$ and the states may be nonorthogonal. The noisy quantum channel can be described by a general quantum evaluation operator $\hat{S}$ with kraussian representation

$$\hat{S}\rho = \sum_\mu A_\mu^\dagger \rho A_\mu, \quad \sum_\mu A_\mu A_\mu^\dagger = \hat{1}. \tag{3}$$

This operators must be linear, completely positive and trace-preserving [1,4]. As follows from definition of quantum information transmission, a possible distortion of entanglement of ρ must be taken into account. In other words definition of mutual quantum information must contain the possible distortion of relative phases of quantum ensemble $\{|\psi_i\rangle\}$. Mutual quantum information is defined as [2]

$$I(\rho; \hat{S}) = S(\hat{S}\rho) - S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)), \tag{4}$$

$$\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)) = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R| \otimes \hat{S}(|\psi_i\rangle\langle\psi_j|). \tag{5}$$

Where $S(\rho)$ is the entropy of von Newman and ψ^R is a purification of ρ

$$|\psi^R\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle \otimes |\phi_i^R\rangle, \quad \langle\phi_j^R|\phi_i^R\rangle = \delta_{ij}, \tag{6}$$

$$tr_R |\psi^R\rangle\langle\psi^R| = \rho, \tag{7}$$

here $\{|\phi_i^R\rangle\}$ is some orthonormal set. The definition is independent from concrete choice of this set [1]. The mutual quantum information is the decrease of the entropy after acting

of $\hat{S}$ due to the possible distortion of entanglement state. This quantity is not symmetric with respect to interchanging of input and output and can be positive, negative or zero in contrast with the Shannon mutual information in classical theory.

It has been shown that (4) can be the upper bound of the capacity of a quantum channel [3,11]. Using this value the authors [3] have been proved the converse coding theorem for quantum source with respect to the entanglement fidelity [1]. Only this fidelity is adequate for quantum data transmission or compression.

In the [2] the authors prove data processing inequality

$$I(\rho; \hat{S}_1) \geq I(\rho; \hat{S}_2 \hat{S}_1). \tag{8}$$

In [3] we found the alternative derivation of this result which is more simple than derivation of [2]. In this paper we show that this equation can be strengthened. Data processing inequality is very important property of mutual information. This is an effective tool for proving general results and the first step toward identification a physical quantity as mutual information.

Now we brief recall the derivation of data processing inequality in the general case. The formalism of relative quantum entropy is very useful in this context [5,7].

Quantum relative entropy between two density matrices ρ_1, ρ_2 is defined as follows

$$S(\rho_1 || \rho_2) = tr(\rho_1 \log \rho_1 - \rho_1 \log \rho_2). \tag{9}$$

This quantity was introduced by Umegaki [6] and characterizes the degree of 'closeness' of density matrices ρ_1, ρ_2. The properties of quantum relative information were reviewed by M.Ohya [5]. Here are mentioned only two basic properties

$$S(\rho_1 || \rho_2) \geq S(\hat{S}\rho_1 || \hat{S}\rho_2). \tag{10}$$

$$S(c\rho_1 + (1-c)\sigma_1 || c\rho_2 + (1-c)\sigma_2) \leq cS(\rho_1 || \sigma_1) + (1-c)S(\rho_2 || \sigma_2). \tag{11}$$

Where $0 \leq c \leq 1$. The first inequality was proved by Lindblad [7]. We have

$$\begin{aligned} & S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|) || \hat{1}^R \otimes \hat{S}(\rho^R \otimes \rho)) \\ = \; & -S(\hat{1}^R \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + S(\hat{S}\rho)). \end{aligned} \tag{12}$$

Here

$$\rho^R = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R| \langle\psi_i|\psi_j\rangle. \tag{13}$$

Now from Lindblad inequality we have

$$\begin{aligned} & S(\hat{1^R} \otimes \hat{S}(|\psi^R\rangle\langle\psi^R|)||\hat{1^R} \otimes \hat{S}(\rho^R \otimes \rho)) \\ \geq \; & S(\hat{1^R} \otimes \hat{S}_1\hat{S}_2(|\psi^R\rangle\langle\psi^R|)||\hat{1^R} \otimes \hat{S}_1\hat{S}_2(\rho^R \otimes \rho)). \end{aligned} \tag{14}$$

From this formula we have (8). From (12) we have

$$I(\rho; c\hat{S}_1 + (1-c)\hat{S}_2) \leq cI(\rho; \hat{S}_1) + (1-c)I(\rho; \hat{S}_2)$$

This theorem have been proved in [11].

Now we use some general theorems for the strengthening the ordinary data processing inequality. In the first we strengthen the Lindblad inequality.

Let us assume in formula (10) that

$$\hat{S} = c\hat{C}_1 + (1-c)\hat{C}_2, \tag{15}$$

where $\hat{C}_1$ is defined by kraussian representation $A_\mu = |\mu\rangle\langle 0|, \quad \langle\mu|\acute{\mu}\rangle = \delta_{\mu\acute{\mu}}, \quad \langle 0|0\rangle = 1, \quad 0 \leq c \leq 1$. In other words for any operator ρ $\hat{C}_1\rho = |0\rangle\langle 0|$. Now from (10), (11) we get

$$\begin{aligned} & S(\hat{S}\rho||\hat{S}\sigma) = S(c\hat{C}_1\rho + (1-c)\hat{C}_2\rho||c\hat{C}_1\sigma + (1-c)\hat{C}_2\sigma) \\ \leq \; & cS(\hat{C}_1\rho||\hat{C}_1\sigma) + (1-c)S(\hat{C}_2\rho||\hat{C}_2\sigma) \leq (1-c)S(\rho||\sigma). \end{aligned} \tag{16}$$

We see that if $\hat{S}$ is represented in the form (15) the ordinary Lindblad inequality can be strengthened.

Now we need some general results from operator theory [10]. Let two hermitian operators A and B have the spectrums $a_1 \leq ... \leq a_n$, $b_1 \leq ... \leq b_n$. For the spectrum $c_1 \leq ... \leq c_n$ of the operator $C = A + B$ we have

$$a_1 + b_k \leq c_k \leq b_k + a_n, \quad b_1 + a_k \leq c_k \leq a_k + b_n. \tag{17}$$

where $k = 1, ..., n$. If

$$\rho\prime = \hat{S}\rho = c\hat{C}_1\rho + (1-c)\hat{C}_2\rho$$
$$= c|0\rangle\langle 0| + (1-c)\sigma, \tag{18}$$

and $\rho_1\prime \leq ... \leq \rho_n\prime$, $\sigma_1 \leq ... \leq \sigma_n$ are the spectrums of $\rho\prime$, σ then we have

$$\rho_1\prime - c \leq \sigma_1(1-c) \leq \min(\rho_1\prime, \rho_n\prime - c),$$
$$\max(\rho_1\prime, \rho_k\prime - c) \leq \sigma_k(1-c) \leq \rho_k\prime, \tag{19}$$

where $k = 2, ..., n$. We define $c(\hat{S}, \rho)$ as the minimal eigenvalue of $\rho\prime$ and $c(\hat{S}) = \min_\rho c(\hat{S}, \rho)$ where minimization is taken by all density matrices for the fixed Hilbert space. With the well known results of operator theory [10] we can write

$$c(\hat{S}) = \min_\rho \min_{\langle\psi|\psi\rangle=1} \langle\psi|\hat{S}\rho|\psi\rangle, \tag{20}$$

where the second minimization is taken by all normal vectors in the Hilbert space. For any density matrix ρ we get to the formula (15) where c is defined in (20) and $\hat{C}_2$ is some general evolution operator. Now from (15, 16, 20) we get the strengthened Lindblad inequality

$$(1-c)S(\rho_1||\rho_2) \geq S(\hat{S}\rho_1||\hat{S}\rho_2). \tag{21}$$

Now we can prove the strengthened data processing inequality. Let in (14) $\hat{S}_2$ is represented in the form (15). From (10-15) we get

$$S(\hat{1}^R \otimes \hat{S}_2\hat{S}_1(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R \otimes \hat{S}_2\hat{S}_1(\rho^R \otimes \rho))$$
$$\leq -(1-c)S(\hat{1}^R \otimes \hat{C}_2\hat{S}_1(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + (1-c)S(\hat{C}_2\hat{S}\rho)). \tag{22}$$

And we have

$$(1-c(\hat{S}_2))I(\rho; \hat{S}_1) \geq I(\rho; \hat{S}_2\hat{S}_1). \tag{23}$$

The equation (20), (23) are our final results. Of course there are many evolution operators $\hat{S}$ with $c(\hat{S}) = 0$ but now we show that our results can be nontrivial because for some simple but physically important case $c(\hat{S})$ is nonzero.

Now we consider the simplest example of noisy quantum channel: Two dimensional, two-Pauli channel [8]

$$A_1 = \sqrt{x}\hat{1}, \quad A_2 = \sqrt{(1-x)/2}\sigma_1, \quad A_3 = -i\sqrt{(1-x)/2}\sigma_2, \quad 0 \leq x \leq 1, \tag{24}$$

where $\hat{1}$, σ_1, σ_2 are the unit matrix and the first and the second Pauli matrices. Any density matrix in two-dimensional Hilbert space can be represented in the Bloch form

$$\rho = (1 + \vec{a}\vec{\sigma})/2, \tag{25}$$

where $\vec{a}$ is a real vector with $|\vec{a}| \leq 1$. Now we have

$$\hat{S_{TP}}((1 + \vec{a}\vec{\sigma})/2) = (1 + \vec{b}\vec{\sigma})/2, \tag{26}$$

where $\vec{b} = (a_1 x, a_2 x, a_3(2x-1))$. After simple calculations we get

$$c(\hat{S_{TP}}) = (1 - |2x-1|)/2. \tag{27}$$

We conclude by reiterating of the main result: The quantum data processing inequality can be strengthened. Such strengthened inequality also exist in classical information theory [9].

References

[1] B.Schumacher, Phys. Rev. A **54**, 2614, (1996).

[2] B.Schumacher and M.A. Nielsen, Phys. Rev. A **54**, 2629, (1996).

[3] A.E. Allahverdyan and D.B. Saakian, eprint quant-ph/9702023.

[4] K.Kraus, Ann. Phys., **64**, 311, (1971).

[5] M.Ohya, Rep. Math. Phys., **27**, 19, (1989).

[6] H.Umegaki, Kodai Math. Sem. Rep., **14**, 59, (1962).

[7] G.Lindblad, Commun. Math. Phys., **40**, 147, (1975).

[8] C.H. Bennett, C.A. Fuchs, J.A. Smolin, eprint quant-ph/9611006.

[9] I.Csiszar, J.Korner, Information Theory, 1982.

[10] F.Gantmacher, The Theory of matrices. Moscow, Nauka 1983.

[11] H.Barnum, M.P. Nielsen, B.Schumacher, eprint quant-ph/9702049.